Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species

Caroline A. A. van Heijningen[a,b,1], Jos de Visser[a], Willem Zuidema[a,c], and Carel ten Cate[a,b]

[a]Behavioral Biology, Institute of Biology Leiden, Leiden University, Sylvius Laboratory, Sylviusweg 72, P.O. Box 9505, 2300 RA, Leiden, The Netherlands; [b]Leiden Institute for Brain and Cognition, Postzone C2-S, P.O. Box 9600, 2300 RC, Leiden, The Netherlands; and [c]Institute for Logic, Language, and Computation, University of Amsterdam, P.O. Box 94242, 1090 GE, Amsterdam, The Netherlands

Edited by Barbara H. Partee, University of Massachusetts, Amherst, MA, and approved October 8, 2009 (received for review July 22, 2009)

According to a controversial hypothesis, a characteristic unique to human language is recursion. Contradicting this hypothesis, it has been claimed that the starling, one of the two animal species tested for this ability to date, is able to distinguish acoustic stimuli based on the presence or absence of a center-embedded recursive structure. In our experiment we show that another songbird species, the zebra finch, can also discriminate between artificial song stimuli with these structures. Zebra finches are able to generalize this discrimination to new songs constructed using novel elements belonging to the same categories, similar to starlings. However, to demonstrate that this is based on the ability to detect the putative recursive structure, it is critical to test whether the birds can also distinguish songs with the same structure consisting of elements belonging to unfamiliar categories. We performed this test and show that seven out of eight zebra finches failed it. This suggests that the acquired discrimination was based on phonetic rather than syntactic generalization. The eighth bird, however, must have used more abstract, structural cues. Nevertheless, further probe testing showed that the results of this bird, as well as those of others, could be explained by simpler rules than recursive ones. Although our study casts doubts on whether the rules used by starlings and zebra finches really provide evidence for the ability to detect recursion as present in "context-free" syntax, it also provides evidence for abstract learning of vocal structure in a songbird.

biolinguistics | discrimination learning | rule learning | syntax

Studying how different features of human language relate to vocal and cognitive abilities in other animals can provide insights into the evolution of the language faculty. A key feature of language is its syntax: the system of rules that govern the construction of sentences by combining smaller elements such as words. It is broadly accepted that the rules that structure animal vocalizations are much simpler than those that structure human language, but it is hard to pinpoint the critical difference. In an influential but controversial paper, Hauser, Chomsky, and Fitch (1) formulated various possible hypotheses, one of them stating that "recursion is the only uniquely human component of the faculty of language" (see also below).

Recursion is typically defined as the embedding of a unit of a particular category inside a larger unit of the same category. In the context of language this enables so called "discrete infinity," whereby an unbounded number of different sentences can be constructed from a limited number of different words. For instance, the sentence "the woman saw," can be embedded in "the man left," which results in: "the man the woman saw left." When we represent the noun phrases with As and the verbs with Bs, this sentence follows an AABB structure, with an AB "recursively" embedded inside another AB. (But see below; the *Discussion* and ref. 2 for the concerns about the actual definition of linguistic recursion; its relation with mathematical recursion and its usage by humans.)

A formalism often used to describe such center-embedded recursion is that of context-free grammars, which form a distinct level in a popular complexity metric known as the Chomsky Hierarchy; the A^nB^nstructure (requiring an arbitrary but equal number of As and Bs) is an example of a structure that cannot be generated by grammars on lower levels on the Chomsky Hierarchy. This is because it requires an unbounded memory to make sure as many Bs are produced as there were As in a string. No such memory is required to produce sentences that follow the $(AB)^n$ structure (such as ABAB or ABABABAB) in which the next AB is concatenated to the previous AB; This is an example of a structure that can be generated by a Finite State Grammar.

It is widely accepted that finite-state grammars are insufficient for describing the syntax of human language (3). In contrast, there is no evidence to date showing that animal vocalizations are also beyond the power of finite-state grammars. However, there is a debate about whether animals have the cognitive ability to detect more complex syntactic rules in artificially constructed stimuli. Until now, two studies have addressed this issue in other species, one on non-human primates and one on a songbird species.

Fitch and Hauser (4) trained and tested humans and cotton-top tamarins (*Saguinus oedipus*) in a familiarization/discrimination paradigm for the ability to distinguish between the above mentioned types of structures. The auditory stimuli consisted of (human) consonant-vowel (CV) syllables spoken by males and females organized in A^nB^n or $(AB)^n$ structures (with $n = 2$ or $n = 3$). While humans were sensitive to a change in structure in either direction, the tamarins did not seem to notice a transition from the A^nB^n to the $(AB)^n$ structure. Fitch and Hauser (4) interpreted this as evidence that tamarins are unable to learn context-free languages and are limited to the less complex, non-hierarchical finite-state languages such as the ABAB structure.

The experiments gave rise to a lively debate, ranging from questions such as whether the stimuli really represented a suitable test of the ability to detect recursive structures, to whether animals as well as humans can "solve" the task by using less complex, non-recursive strategies (2, 5–12).

From a methodological perspective, one concern about the study is that non-conspecific, human speech stimuli were used to test the tamarins, while this species (like most primate species) also has a limited vocal repertoire and lacks the ability for vocal learning. The failure of the tamarins might thus be explained by more basic features of their perceptual and learning abilities rather than the ability to represent and process syntactic structure. A stronger test for the human uniqueness of syntax comes from comparative studies with animals, such as songbirds, showing vocal learning and complex vocalizations, in particular from studies where the stimuli consist of units present in their natural vocalizations. Such an

Author contributions: C.A.A.v.H., W.Z., and C.t.C. designed research; C.A.A.v.H. and J.d.V. performed research; C.A.A.v.H., J.d.V., and W.Z. analyzed data; and C.A.A.v.H., W.Z., and C.t.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: c.a.a.van.heijningen@biology.leidenuniv.nl.

This article contains supporting information online at www.pnas.org/cgi/content/full/0908113106/DCSupplemental.

experiment was carried out in the starling (*Sturnus vulgaris*) (13). Starlings are songbirds that are known for their complex learned song that consists of several "song phrase types" that are repeated two or more times before the next phrase type is sung (14). The grammar stimuli were constructed from two of these natural phrase types: "rattles" and "warbles."

Gentner et al. (13) trained starlings in a go/no-go experiment, a different paradigm than the habituation/dishabituation experiment that was used to test discriminatory abilities in tamarins. In this go/no-go experiment, the bird has to respond to one type of stimuli, but withhold responses to the other. Contrary to the tamarin experiment by Hauser and Fitch (4) in which recognizing each grammar was tested separately by presenting violations (the other grammar) in a habituation/dishabituation experiment, in this go/no-go experiment, the discrimination between the two grammars was tested directly. The authors reported that starlings, unlike the cotton-top tamarins, were able to "accurately recognize (...) a recursive, self-embedding, context-free grammar," "classify new patterns defined by the grammar and reliably exclude agrammatical patterns." This claim entails that the capacity to detect recursion in auditory sequences would not be uniquely human. In support of the claim, the study included a range of controls for alternative (simpler, but still abstract) rules that the birds could have used to distinguish the two stimuli sets.

Like the tamarin experiment, the starling experiment by Gentner et al. (13) received a (similar) variety of comments that were brought up before in response to HCF (1) and Fitch and Hauser (4). These mainly concerned the actual recursiveness of the test stimuli and the way they were presented; the level of abstract rule learning that the starlings showed; the relatively high number of trials the starlings needed to learn the task and the correctness of the implications the authors mention for the unique characteristics in human language (2, 9–12, 15, 16).

But while the discussion on the experiment in terms of whether it is an appropriate test of the presence of recursion lingers on, a more basic question about what the starlings have actually learned has received little attention (but see refs. 2 and 16): did they learn an abstract rule allowing them to generalize from familiar to novel stimuli? Gentner et al. (13) showed that starlings generalized from familiar rattle and warble structures to unfamiliar ones, but not that they generalized to other phrase types as well. Therefore, the birds could also have based their discrimination on a phonetic categorization, restricted to rattle and warble phrase types. That is, they might classify all "rattle-warble-rattle-warble" sequences as one type, different from "rattle-rattle-warble-warble" sequences, based on the overall acoustic similarity alone [and with the appropriate similarity metric, this would generalize to larger n without a need for a syntactic analysis (17)]. So the question whether the birds have the abstract computational abilities that enable them to discriminate between the two stimuli sets strictly based on order information is still open.

To test this, the birds should be presented with a new set of song stimuli, consisting of new phrase types (for instance structures constructed of Cs and Ds). In this case, the birds cannot use phonetic generalization, but have to attend to the sequence of phrase types, requiring some level of abstract computational skill. In the current paper, we present such an experiment with another songbird species, the zebra finch (*Taeniopygia guttata*). Using a similar setup as the Gentner et al. study (13)—including, importantly, the use of a probe methodology to test for a range of alternative hypotheses—we extend their results by explicitly testing for generalization and by a more in-depth statistical analysis of the probes.

The zebra finch is widely used as a model species in neuro- and behavioral sciences and knowledge about the perceptual abilities of this species would be greatly beneficial to these research areas. An additional reason for choosing the zebra finch is the suggestion that the complex perceptual ability in starlings could be related to the relatively complex structure of their natural songs (18). The song of the zebra finch is relatively simple, consisting of a limited number of elements and a relatively stereotyped sequence (e.g., refs. 19–23). Nevertheless, zebra finches are able to discriminate between songs with subtle differences such as the number of repeated elements and the relative position of a certain element in a song (24). So our experiment also tests whether birds with much simpler songs are able to detect more abstract patterns in vocal structures.

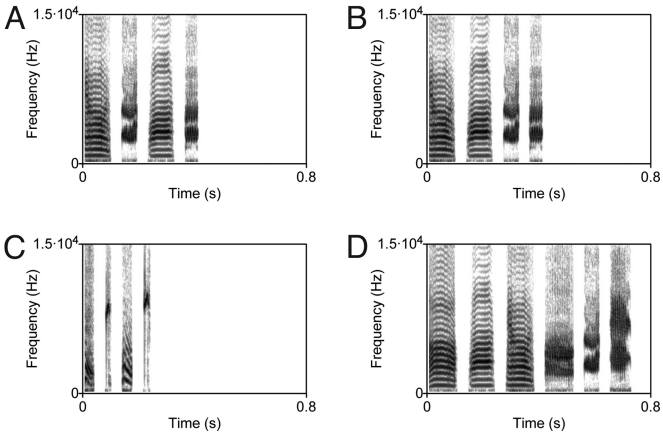

Fig. 1. Spectrograms of constructed song stimuli. (*A*) Stacks and trills in ABAB order. (*B*) Stacks and trills in AABB order. (*C*) Slides and highs in ABAB order. (*D*) Stacks and trills in AAABBB order. Note that the different panels show different exemplars of the different elements used. In particular, *D* illustrates within and between element type variation in the stimulus sets.

The results from this experiment show that zebra finches, like starlings, can discriminate between sets of two syntactically different song stimuli and can generalize this discrimination to unfamiliar exemplars of the same stimulus types.

In addition, when confronted with novel stimulus types arranged identically to the earlier ones, one out of eight birds was able to correctly classify the stimuli, an ability that so far has not been shown in any other songbird. However, additional "probe tests" to examine the rules that the birds used to discriminate between syntactically different sets, suggest that all birds used simpler rules to discriminate between the sets with and without "recursion."

Results

Transfer from One Song to Five Songs of Each Structure. During the training phase, the birds learned to discriminate between a single ABAB and one AABB stimulus (see Fig. 1 *A* and *B*. See *SI Text* for two example song stimuli used in this experiment). As the two structures consisted of identical elements, the only way to discriminate between the two is by learning about the order of the individual elements within the structures, an ability demonstrated in zebra finches before (24).

The simplest mechanism to do this would be to learn the sound characteristics of each structure; that is, by rote memorization. We tested this by presenting additional stimuli (2×4) of the same structures but constructed of different element exemplars. If the birds based their discrimination on the individual sound characteristics of the two stimuli, they would perceive this as a new task and show a considerable drop in DR after transfer to $\approx$0.50 and a d′ of around zero, meaning no discrimination (see refs. 13 and 25 for method of calculating the d′ and its confidence interval). Alternatively, if the birds learned a more general rule either related to the sequential structure of the sounds, or their phonetic structure, they should be able to generalize to the

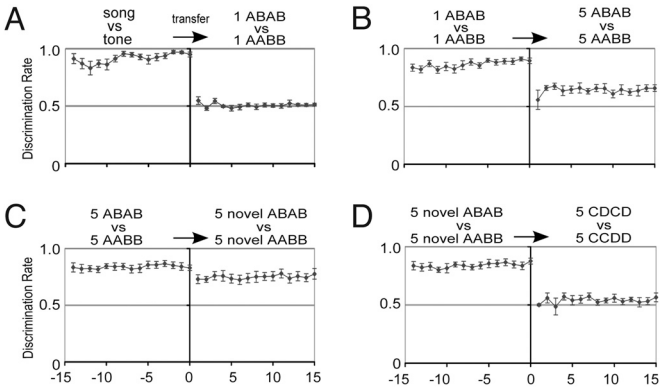

Fig. 2. Average discrimination ratios (DRs) per blocks of 100 trials before and after transfer to new stimuli (± SEM.).

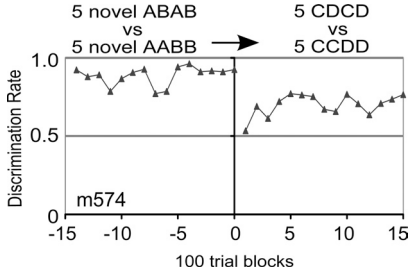

Fig. 3. Discrimination ratios per blocks of 100 trials of bird m 574 before and after transfer to song stimuli of new element types.

additional stimuli and show a DR >0.50 (and a d′ higher than zero). Fig. 2*B* shows the average discrimination ratios per 100 trials before and after transfer to 5 ABABs and 5 AABBs. Directly after transfer, there is an initial drop of the DR to 0.56, which is slightly higher than random performance, but the birds learned to discriminate faster than during the previous task. This can be concluded from the d' measures: for the first two 100-trial blocks, six out of eight birds discriminated significantly between the two sets, and seven out of eight were able to do so during blocks 3–5 after transfer (see Table 1). Hence, we can conclude that most birds were able to generalize to the additional stimuli.

Transfer from Five Songs of Each Stimulus Type to Five Novel Songs of Each Type. With additional training, all birds reached criterion performance. However, the question remains whether the birds learned an abstract computational rule to discriminate between these ABAB songs and AABB songs or whether there is a lower-level explanation (i.e., learning the 10 songs by rote memorization or by phonetic generalization).

To test whether the zebra finches are able to generalize independent of the characteristics of the individual elements in the stimuli, a different set of stimuli of the same element categories was introduced. From stimuli constructed of a_{1-5} and b_{1-5}, each bird was transferred abruptly to stimuli constructed from a_{6-10} and b_{6-10} when it reached criterion performance. See Fig. 2*C* for average discrimination ratios per 100 trials before and after transfer to 2 × 5 novel songs of each structure. After transfer, the average DR remained >0.50. The slight drop indicates that the birds do notice a difference between the first and second sets, but they treated the novel stimuli similar to the previous ones. Seven out of eight birds generalized to the new stimuli within the first two blocks and all birds did so within blocks 3–5 (see Table 1). This outcome suggests that they used a general rule instead of rote memorization of all individual songs to base discrimination on.

It is here that the question arises what exactly the birds use to generalize: an abstract rule (as suggested by Gentner et al., ref. 13), or a lower-level process, such as phonetic generalization. The use of an abstract computational rule would predict a transfer of the discrimination to songs of any element type, as long as they occur in a same order. The question thus is whether the birds can generalize independent of element type.

Transfer from Five Songs of Each Stimulus Type to Five Songs with New Element Types (As and Bs to Cs and Ds). To test whether the zebra finches are able to generalize to new songs independent of element type, each bird was transferred abruptly to Cs and Ds; that is, other types of elements instead of As and Bs, when it reached criterion performance (see Fig. 1*C*). The average DRs before and after transfer are shown in Fig. 2*D*. The average DR after transfer stayed just >0.50, but all d″s for the initial two blocks after transfer were not significant. For blocks 3–5 however, three of the eight birds showed significant discrimination (see Table 1). However, two of these three birds dropped to chance performance again (DR of 0.50) shortly after this interval. But the third bird, m574, remained on a high discrimination level (see Fig. 3). Only after training did the seven other birds learn to discriminate between the new stimuli sets, indicating that they perceived this as a new task.

So we conclude that seven out of eight birds were unable to generalize the discrimination of the stimuli to novel ones using structure cues only. However, there was one zebra finch that showed a clearly different response pattern and was likely to be able to generalize independent of element types and therefore used a more abstract general rule.

Testing for the Use of Other, Simpler Rules to Discriminate Between ABABs and AABBs. One bird (m574) showed that he was capable of generalizing across element types and so apparently used a more abstract rule for distinguishing between the stimuli sets than the other birds. However, the rule it used might not have been "recursive," but could have been a more simple, lower-level strategy. For instance, ABAB and AABB differ in their first two elements and therefore can be discriminated based on this difference instead of all four elements ("primacy rule"). Similarly, discrimination might be based on the last two elements only ("recency rule").

Table 1. Average d′s after each transfer (± SEM)

Transfer		Blocks 1 and 2				Blocks 3 to 5			
From	To	d′	SEM	Min L bound CI	max L bound CI	d′	SEM	min L bound CI	max L bound CI
Song/tone	1ABAB/AABB	0.04	0.09	−0.54	0.21	0.07	0.06	−0.53	−0.01
1ABAB/AABB	5ABAB/AABB	0.67	0.13	−0.33	0.83	0.68	0.16	−0.25	1.08
5ABAB/AABB	2× 5 novel	1.10	0.16	−0.09	1.29	1.31	0.17	0.32	1.67
2× 5 novel	2× 5 new element types	0.09	0.07	−0.47	−0.03	0.25	0.10	−0.47	0.40

Min L bound CI: lowest lower bound of the confidence interval (of individual birds). Max L bound CI: highest lower bound of the confidence interval (of individual birds).

Table 2. Average response to the two consecutive blocks of 30 probes of each probe type

		Probe type								
Bird ID	S+	ABAB	AABB	AAAB	AAAA	BBBB	BAAB	ABBA	AAABBB	AAAABBBB
570	AABB	1.5	28.0	22.5	22.5	0.0	8.5	1.5	17.5	14.0
574	AABB	1.0	24.5	8.0	8.0	23.0	9.5	1.5	20.0	21.0
583	ABAB	22.0	0.0	0.0	0.0	25.0	2.5	26.0	0.0	0.5
602	ABAB	25.0	3.0	4.5	2.0	22.0	1.0	26.5	1.0	1.0

These lower-level rules were tested in a probe testing phase during which several probes were presented, similar in design to the probes used in Gentner et al. (13). In this testing phase, four birds were exposed to probe songs. Probe songs were not reinforced to avoid additional learning and presented in only 20% of all stimulus presentations to prevent extinction of the pecking behavior. The other 80% of stimulus presentations were the reinforced stimulus songs to which the birds were exposed during the previous phase (Table 2 provides individual responses to these probes; see Fig. 1*D*, an example.)

To test whether the birds showed a differential response to the probes, we tested for each bird individually whether the responses to the probes differed, which it did (binomial test in R, y~probetype, family = quasibinomial). All deletion *P* values for probe type were significant (largest *P* value was <0.05), indicating that all four birds did not respond in a random manner to the probe songs.

However, the pattern of responses to probes did not fit either of the grammars used to define the training stimuli. Like Gentner et al. (13), we considered the alternative hypotheses that for distinguishing the training stimuli the birds used a strategy based only on the first transition in each sequence (AA vs. AB, "primacy"), the last transition (BB vs. AB, "recency") or any transition (the presence vs. absence of AA, BB or BA, "bigram"). Gentner et al. (13) try to exclude these alternative strategies by showing that the birds are significantly better at distinguishing between the two sets of training stimuli than between a contrasting pair of probe stimuli. That is, the d′ between ABAB and AABB is significantly larger than between BAAB and ABBA, which Gentner et al. (13) take as excluding the bigram hypotheses.

However, given that the d′ is significantly smaller for all pairs of probe stimuli, including the "context-free" probes AAABBB and AAAABBBB, that is not the relevant test. We performed a maximum likelihood analysis instead, evaluating for every individual bird which of the considered alternative hypotheses best explains the observed data. From such an analysis (as shown in Table 2) we deduced that m574, the bird that was able to generalize to new element types, did so without using the context-free strategy. If he had acquired this strategy, he should not respond to the following probes, AAAB, AAAA, BBBB, BAAB and ABBA. The results show that he did in fact respond to for instance, BBBB. His pattern of response best fits a BB-recency strategy (responding selectively to sequences ending in BB). Two other birds most likely used "bigram AA" to distinguish between the sets; one bird used "primacy AA."

Discussion

This paper addressed the question whether songbirds are able to discriminate between vocal structures with and without a "center-embedded recursive" structure. The results of our experiment show that all zebra finches learned to discriminate between two stimulus sets that solely differed in the order of the elements. This shows that a songbird species lacking complex and variable songs can be highly sensitive to song structure, corroborating results in a different context by Verzijden et al. (24).

We show that zebra finches are able to generalize from training exemplars to new song stimuli consisting of new elements of the same element types. A number of birds showed a slightly delayed generalization response after certain transfers to new stimuli. The magnitude of their discrimination (described by the DR and d′) was reduced directly after transfer (blocks 1 and 2), which would not be expected for generalization. However, their recovery during blocks 3–5 was faster than would be expected if they had perceived this as a new task (i.e., compared to for instance, the response behavior after transfer to the first ABAB and AABB).

Until this stage, the experiment ran parallel to Gentner et al. (13). However, our main objective was to test whether songbirds are able to generalize independently of phonetic characteristics of the element types that are used in the stimulus songs. The results of our additional transfer stage showed that only one bird was able to discriminate the stimuli independent of element type. This zebra finch was therefore able to use more abstract skills than the others. However, additional testing showed that, just as other zebra finches, this bird did not use a "recursive" strategy, but, a simpler strategy (although still abstracting over phonetic detail). From an efficiency point of view, this makes sense as it would be inefficient to use relatively complex rules, when simpler rules suffice. We conclude that the current experimental paradigm, although yielding interesting results, has not answered the questions whether zebra finches have the ability to detect recursion.

However, this conclusion also applies to previous studies on this issue. For the starling experiment of Gentner et al. (13) for instance, it is still not clear whether the data allow the rejection of the primacy rule, for two reasons.

First, the mean d′ between AAAA and ABBA was significantly lower than between AABB and ABAB, but it was still significantly >0 (0.99 ± 0.40 SEM). Consequently, it cannot be ruled out that the starlings (at least partly) used a primacy rule instead of a "recursion" rule.

Second, our zebra finches used different individual strategies (see Table 2) to discriminate between the stimulus songs. Gentner et al. (13) reported the mean d′ ± SEM instead of the individual d″s. Therefore, if starlings also show individually different strategies, every other comparison than between AABB and ABAB is likely to produce a less clear result, as all other comparisons are a mixture of one or more birds that did or did not use the tested strategy. This would mean that contrary to the authors' conclusions, the evidence whether starlings are able to detect recursion is inconclusive.

Interestingly, the same uncertainty remains about human abilities in artificial language learning tasks. The findings in (4) lacked the appropriate controls, but later replications claim that the AABB/ABAB task, which Fitch and Hauser (4) designed to obtain evidence for recursion, can be solved by humans using a simpler strategy instead (5, 9) or by a conscious counting strategy that seems unrelated to language (11). At present, there is thus no convincing demonstration of the use of recursive rules in artificial language learning in any species. It remains a challenge to design experiments on artificial rule learning and its underlying mechanisms that unambiguously exclude simpler explana-

tions for discrimination between training structures, a problem that also faces experiments in humans (e.g., ref. 26). Combined with the fact that it is far from clear that recursion actually is the only unique aspect of human language (see refs. 27 and 28), or even a particularly relevant feature (29), it might be time to switch the attention to the pattern recognition abilities of zebra finches and other species, which are remarkable in their own right, quite independently from the question of the origin of recursion in human language. With this study, we have uncovered some details of those abilities in zebra finches, extended the experimental paradigm to address such questions and hopefully contributed to putting the claim for or against the human uniqueness of recursion in the right perspective.

Materials and Methods

Subjects and Housing. Eight zebra finches from our breeding colony (six males and two females; age range: 137–363 days at the start of the experiment) were individually trained and tested in this experiment. They had no previous experience with similar experiments.

Before the experiment, the birds were housed in the breeding colony in same sex groups of two to six individuals and on a 13.5 L:10.5 D schedule at 20–22 °C. Drinking water, cuttlebone and a commercial tropical seed mixture enriched with minerals was available ad libitum. Twice a week, the birds received some egg food and seedlings from the seed mixture.

Apparatus. During the experiment, subjects were individually housed in an operant conditioning chamber [70 (l) × 30 (d) × 45 (h) cm], constructed of wire mesh front and side walls, and a ply wood back wall. The cage was placed in a sound-attenuated chamber. Water and cuttlebone were available ad libitum. A fluorescent tube (Lumilux De Luxe Daylight, Osram) was placed on top of the cage, with the same light/dark schedule as the breeding colony (except during punishment reinforcement, see go/no-go procedure). The back wall contained two red pecking sensors with red LED lights and a food hatch. Stimuli were played through a speaker (Vifa MG10SD09–08) located 1 m above the operant conditioning chamber and calibrated to a (peak) output of approximately 70 dB (SPL meter, RION NL 15, RION) at the food hatch. The fluorescent tube, the two sensors, the food hatch, and speaker were connected to a custom-made operant conditioning chamber controller that also registered the sensor pecks. To monitor their condition, the birds were weighed before and after the experiment and their food intake was measured daily. The study was conducted according to Association for the Study of Animal Behavior guidelines on animal experimentation as well as to the Dutch law on animal experimentation. The Leiden Committee for animal experimentation (DEC) approved the experiment under number 06150.

Song Stimuli Construction. We constructed stimulus songs, each containing four units, as in the earlier studies (4, 13). In our case, these units were elements obtained from natural zebra finch song from our zebra finch song database (consisting of undirected song, for recording specifications see ref. 23).

Four element types occurring in natural zebra finch song (''flats,'' ''slides,'' ''highs,'' and ''trills'' (see Fig. 1) were selected that were easily distinguishable by visual inspection. The first three have previously been recognized as element types by several authors (e.g., 19–22). The fourth, the trill, characterized by rapid and repeated frequency modulation and, is less well known but occurs in songs from our laboratory (23) and has been previously described as a ''click'' (30) and as a ''buzz'' (31).

An element library of 40 (10 × 4) elements was constructed by selecting elements from songs in our zebra finch song lab database. All elements (as well as the shaping stimuli) were ramped (3 ms) and rms (0.1) equalized in PRAAT (version 4.5.08, www.praat.org).

Subsequently, 40 stimulus songs were constructed, each consisting of four different elements, with 40 ms as an inter-element pause. Each four-element song consisted of two element types, As and Bs, and were ordered in either the ABAB structure or the AABB structure. In a later phase we used elements of another type (Cs and Ds). Songs always consisted of non-identical elements within element types. So $a_1b_1a_2b_2$ was a member of the ABAB type set and $a_1a_2b_1b_2$ for its equivalent in the AABB type set. For more information on probe stimuli construction, see *SI Text*.

To pretrain the birds for the experiment, two other stimuli were constructed which were similarly ramped and equalized as the elements from the element database: a natural zebra finch song motif, 0.58 s in duration, randomly chosen from our zebra finch song database and a 2-kHz tone of identical duration as the S^+, also constructed in PRAAT.

To control for order effects, the element types of the stimuli songs that the experiment was started with (As and Bs) was counterbalanced between birds. So the other half of the birds started with Cs and Ds and were in this phase transferred to As and Bs.

Go/No-Go Procedure. The birds were trained in a go/no-go procedure with a food reward to test their ability to discriminate between stimuli of the ABAB and AABB structures.

A peck on the left sensor, when its LED was on, initiated a playback of a positive ''go'' stimulus song (S^+, the natural zebra finch song motif). After this song was finished, the LED in the right sensor was switched on and pecking on it opened the food hatch for 10 s (positive reinforcement). When the bird pecked at the right sensor within 6-s response interval for at least 75%, the negative ''no-go'' stimulus (S^-, the tone), was introduced in 60% of the cases the bird pecked the left sensor. Subsequently, the bird had to learn not to peck at the right sensor after hearing an S^- stimulus, or the lights would go out for 15 s (negative reinforcement).

After the bird reached criterion performance (>75% response after S^+ and <25% response after S^- for at least two full consecutive days) and a phase where the ratio of S^+/S^- was set to 50/50 (in random order), the training with the two stimuli sets began. The type of stimulus, either ABAB or AABB, was mapped on the S^+ and S^-, and was counterbalanced between birds. Half of the birds received the ABAB songs as the S^+ stimuli and the AABB songs as the S^- stimuli.

Statistics. Discrimination by each zebra finch was visualized by plotting the DR (discrimination ratio, calculated as the percentage response to S^+ divided by the sum of the percentage response to S^+ and the percentage response to S^-), but was measured with d′ (both per 100 trials), which is independent of response bias (25). A Confidence Interval for d′ with a lower bound >0 was considered as significant discrimination between two stimuli sets. Two intervals after transfer to a new set of stimuli were tested: the first two blocks and blocks 3–5. See refs. 13 and 25 for method of calculating the d′ and its confidence interval. Some birds showed a continued response bias during one or more phases. If this occurred in combination with a very low S^-, and hence a high DR we decided to transfer the particular bird to the next phase when their DR was >0.75 for at least two full consecutive days (alternative learning criterion).

For the statistics on the data from the probe testing, see *SI Text*.

Training Phase: Transfer from Song and Tone to One ABAB and One AABB Song. After the pretraining with a song and a tone and the birds reached criterion performance, the stimuli were abruptly changed (transferred) to one exemplar of ABAB and one of AABB. Just before transfer, the birds ($n = 8$) had an average d′ of 3.49 ± 0.26 (SEM) and a range for the lower bound of the 95% confidence interval (CI range) of 1.95–3.45 calculated for the last two blocks of 100 trials each.

It can be seen in the graph (Fig. 2*A*, DR ≈0.50) and from the d′ value and its CI range in Table 1 (d′ 0.04 ± 0.09, lower bounds of the individual CIs between −0.54 and −0.21) that this is indeed the case for the first two blocks of 100 trials after transfer (d′ is near zero, which means no discrimination). The birds seem to start from scratch again, that is, there was no obvious transfer of the go/no-go paradigm to novel stimuli, as indicated by the finding that none of the birds was significantly discriminating between the sets during blocks 3–5 after transfer.

One bird did show significant discrimination in these first two 100 trial blocks, but thereafter his d′ dropped <0 and was no longer significant during blocks 3–5. However, after training, all birds discriminated the AABB and ABAB stimuli on or above criterion level.

ACKNOWLEDGMENTS. We thank Erwin Ripmeester and Tom van Dooren for help with the statistics; Rinus Heijmans and Frits van Tol for constructing the operant conditioning chambers; Rob van der Linden and Ap Gluvers for developing the electronics and software for the operant conditioning chamber controllers; and two anonymous referees for valuable comments on the manuscript. This work was supported by Netherlands Organization for Scientific Research Grant 051.07.002 (to W.Z.).

1. Hauser MD, Chomsky N, Fitch WT (2002) The faculty of language: What is it, who has it, and how did it evolve? *Science* 298:1569–1579.
2. Corballis MC (2007) Recursion, language and starlings. *Cogn Sci* 31:697–704.
3. Chomsky N (1957) *Syntactic Structures* (Mouton, The Hague).
4. Fitch WT, Hauser MD (2004) Computational constraints on syntactic processing in a nonhuman primate. *Science* 303:377–380.

5. Perruchet P, Rey A (2005) Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? *Psychon Bull Rev* 12:307–313.
6. Zuidema WH (2005) in *The Major Transitions in the Evolution of Language, Theoretical and Applied Linguistics* (dissertation) (University of Edinburgh, Edinburgh).
7. Kochanski G (2004) Comment on ''Computational constraints on syntactic processing in a nonhuman primate.'' http://kochanski.org/gpk/papers/2004/FitchHauser/FitchHauserScienceLetter.pdf
8. Bahlmann J, Gunter TC, Friederici AD (2006) Hierarchical and linear sequence processing: An electrophysiological exploration of two different grammar types. *J Cogn Neurosci* 18:1829–1842.
9. De Vries MH, Monaghan P, Knecht S, Zwitserlood P (2008) Syntactic structure and artificial grammar learning: The learnability of embedded hierarchical structures. *Cognition* 107:763–774.
10. Corballis MC (2006) On phrase structure and brain responses: A comment on Bahlmann, Gunter, and Friederici. *J Cogn Neurosci* 19:1581–1583.
11. Hochmann JR, Azadpour M, Mehler J (2008) Do humans really learn A^nB^n artificial grammars from exemplars? *Cogn Sci* 32:1021–1036.
12. Liberman M (2006) Language log on starlings. http://itre.cis.upenn.edu/~myl/languagelog/archives/003076.html
13. Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC (2006) Recursive syntactic pattern learning by songbirds. *Nature* 440:1204–1207.
14. Eens M (1997) Understanding the complex song of the European starling: An integrated approach. *Adv Study Behav* 26:355–434.
15. Premack D (2007) Human and animal cognition: Continuity and discontinuity. *Proc Natl Acad Sci USA* 104:13861–13867.
16. Marcus GF (2006) Startling starlings. *Nature* 440:1117–1118.
17. Suhara Y, Sakurai A (2007) A simple computational model for classifying small string sets, International Congress Series 1301, Elsevier: 270–273. doi: 10.1016/j.ics.2006.11.006.
18. Yip M (2006) The search for phonology in other species. *Trends Cogn Sci* 10:442–446.
19. Price PH (1979) Developmental determinants of structure in zebra finches. *J Comp Physiol Psychol* 93:260–277.
20. Scharff C, Nottebohm F (1991) A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system—implications for vocal learning. *J Neurosci* 11:2896–2913.
21. Zann R (1993) Variation of song structure within and among populations of Australian zebra finches. *Auk* 110:716–726.
22. Sturdy CB, Phillmore LS, Weisman RG (1999) Note types, harmonic structure, and note order in the songs of zebra finches (*Taeniopygia guttata*). *J Comp Psychol* 113:194–203.
23. Holveck MJ, Viera de Castro ACV, Lachlan RF, ten Cate C, Riebel K (2008) Accuracy of song syntax learning and singing consistency signal early condition in zebra finches. *Behav Ecol* 19:1267–1281.
24. Verzijden MN, Etman E, van Heijningen C, van der Linden M, ten Cate C (2007) Song discrimination learning in zebra finches induces highly divergent responses to novel songs. *Proc R Soc London B* 274:295–301.
25. MacMillan NA, Creelman CD (2005) *Detection Theory: A User's Guide* (Lawrence Erlbaum Associates Mahwah, NJ), 2nd Ed.
26. Seidenberg MS, MacDonald MC, Saffran JR (2002) Does grammar start where statistics stop? *Science* 298:553–554.
27. Christiansen MH, Kirby S (2003) Language Evolution (Oxford Univ Press, Oxford).
28. Pinker S, Jackendoff R (2005) The faculty of language: What's special about it? *Cognition* 97:201–236.
29. Verhagen A (2009) What do you think is the proper location of recursion? Conceptual and empirical issues. *Ling Rev* In press.
30. Williams H, Staples K (1992) Syllable chunking in zebra finch (*Taeniopygia guttata*) song. *J Comp Psychol* 106:278–286.
31. Leadbeater E, Goller F, Riebel K (2005) Unusual phonation, covarying song characteristics and song preferences in female zebra finches. *Anim Behav* 70:909–919.